LOSSY COMPRESSION OF STOCHASTIC HALFTONES WITH JBIG2

Magesh Valliappan and Brian L. Evans

Dept. of Electrical and Computer Eng.
The University of Texas at Austin
Austin, TX 78712-1084 USA
{magesh,bevans}@ece.utexas.edu

Dave A. D. Tompkins and Faouzi Kossentini

Dept. of Electrical and Computer Eng.
University of British Columbia
Vancouver, B.C. V6T 1Z4 Canada
{davet,faouzi}@ece.ubc.ca

ABSTRACT

The JBIG2 standard supports lossless and lossy coding models for text, halftone, and generic regions in bi-level images. For the JBIG2 lossy halftone compression mode, halftones are descreened before encoding. Previous JBIG2 descreening implementations produce high-quality images for clustered dot halftones at high compression rates but significantly degrade the image quality for stochastic halftones, even at much lower rates. In this paper, we develop (1) a flexible, computationally efficient, JBIG2-compliant method for compressing stochastic halftones that reduces noise, artifacts, and blurring; (2) quality measures for linear and nonlinear distortion in compressed halftones; and (3) rate-distortion tradeoffs for the encoder parameters.

1. INTRODUCTION

Digital halftoning converts a continuous-tone image into a bi-level image (halftone) for printing and display on binary devices. Bi-level images consist of a single rectangular bit plane. Pixels are either assigned black or white to create an illusion of continuous shades of gray.

Halftoning by ordered dithering thresholds a grayscale image by using a periodic mask of threshold values. In clustered dot ordered dithering, black dots are clustered in large blobs. Since clustered dot halftones are resistant to ink spread, they are the most common among printed halftones. Stochastic halftoning varies the thresholding according to local statistics in the image in order to shape the quantization noise into the high frequencies where the human visual system is less sensitive. Stochastic halftoning requires more computation and generally yields better visual quality than ordered dithering [1].

Current fax machines only support one *lossless* compression mode that has been optimized for textual data [2] and causes data expansion when applied to halftones. The Joint Bi-Level Experts Group (JBIG) is a subcommittee of both the ISO/IEC and the ITU-T that is developing a second international standard for bi-level image compression for use in printers, fax machines, scanners, and document storage and archiving. JBIG2 adds *lossy* compression and supports several different coding models for text, halftone and generic regions. JBIG2 should be finalized in Fall 1999.

JBIG2 specifies the bit stream syntax, which places strict requirements on decoder designs but leaves much flexibility for encoder designs. An encoder could segment bi-level images into text, halftone, and generic regions, and encode each region separately [3, 4]. For halftone regions, JBIG2 supports very high lossy compression rates. This paper develops a new high-quality lossy JBIG2 encoding method tailored for stochastic halftones. We develop visual quality measures, which we use for rate-distortion tradeoffs. We achieve an additional 35% compression over the best reported method [5]. We implement the method in the JBIG2 codec available at *http://spmg.ece.ubc.ca/jbig2.*

2. BACKGROUND

JBIG2 uses a form of vector quantization for compressing halftones. First, the encoder descreens (inverse halftone) the bi-level image into a grayscale image with reduced spatial resolution. Gray levels serve as indices into a halftone pattern dictionary, which is chosen by the encoder. Then, the encoder encodes the halftone pattern dictionary and grayscale image bitplanes losslessly using the generic mode of operation, and chooses the orientation for the halftoning grid. The decoder first decodes the bitplanes to construct the grayscale image and then constructs the bi-level image by placing the dictionary patterns corresponding to the grayscale values at their appropriate positions and orientations. Thus, the spatial resolution and number of gray levels directly affect the quality and compression ratio. Perceptually lossless compression is achieved by preserving the local average gray level but not the bi-level image itself.

One descreening method [5] maps each non-overlapping $M \times M$ window of halftone pixels to one grayscale pixel value that is equal to the number of black pixels in the halftone window. For non-angled grids, this method yields $M^2 + 1$ gray levels. For angled grids, a mask is used with the window so that adjacent patterns do not overlap, and the number of gray levels is *area of the mask* $+1 \leq M^2 + 1$. This method is conceptually and computationally simple, and works well for ordered dithered halftones. When this method is applied to stochastic halftones, the grayscale image suffers from noise, blur, and artifacts, which degrade the reconstructed halftone quality [5].

For stochastic halftones, inverse halftoning methods in descending order of quality are set theoretic [10], nonlinear denoising [7], adaptive smoothing [6], overcomplete wavelet expansion [8], wavelet denoising [9], projection onto convex sets [11], and vector quantization [12]. The inverse halftones have the same spatial resolution as the halftone but with 6–8 bits of precision per pixel. For higher compression, a JBIG2 encoder would create a grayscale image at lower spatial resolution and fewer gray levels.

M. Valliappan and B. Evans were supported by a US National Science Foundation CAREER Award under grant MIP-9701707. D. Tompkins and F. Kossentini were supported by the Natural Sciences and Engineering Research Council of Canada.

3. PROPOSED ENCODING METHOD

The proposed encoder in Fig. 1 takes a stochastic halftone as input and generates a JBIG2-compliant bitstream as output. The free parameters are grid size and orientation, number of quantization levels, and sharpening control.

The prefilter should suppress high-frequency noise, spurious tones, and Nyquist frequencies in stochastic halftones, and it should have a flat passband response to minimize distortion [13]. To meet these criteria while maintaining computational simplicity, we design a symmetric 3×3 finite impulse response filter with power-of-two coefficients [13]: $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$. The prefilter produces a grayscale image at the same spatial resolution as the halftone.

The decimator consists of a lowpass anti-aliasing filter followed by downsampling by M in each dimension. The filter sums the pixel intensities within the pattern mask. For unangled grids, this corresponds to a cutoff frequency of $\frac{\pi}{M}$ in each dimension. This process is very similar to [5] except that the mask and the window are applied to the filtered grayscale values instead of the bi-level image and that the number of resulting gray levels is significantly higher. In our implementation, we combine the prefilter and the anti-aliasing filter. Since the input would be a bi-level image, we replace multiplications with additions. For a 3×3 smoothing filter and a window size of $M \times M$, we need at most $(M+2)^2$ additions for each grayscale value.

The quantizer uses N gray levels ($N \leq M^2 + 1$). Conventional quantization with 32 or fewer levels ($M \leq 5$) would create false contours [14]. To avoid contouring, we dither the input image using a multi-level version [15] of modified error diffusion [16]. Error diffusion, which is a type of 2-D sigma-delta modulation, shapes the quantization noise into the higher frequencies. Modified error diffusion controls the amount of linear distortion by means of a sharpening control parameter L, which requires an extra addition and multiplication per input pixel. We choose L to compensate for the blurring in the previous stages. To achieve higher compression, we could further quantize the gray levels with a slight loss in quality. In our implementation, we use a four-tap Floyd-Steinberg error diffusion filter which has dyadic coefficients. It is the smallest error diffusion filter known to produce high-quality halftones [13, 17].

The size, shape, and orientation of the patterns directly affect the reconstructed halftone quality. Halftone patterns oriented at 45° can yield perceptually better results [1]. The possible patterns depend on the rendering device. We use methods similar to [5] to generate halftone patterns for both angled and non-angled clustered dot masks.

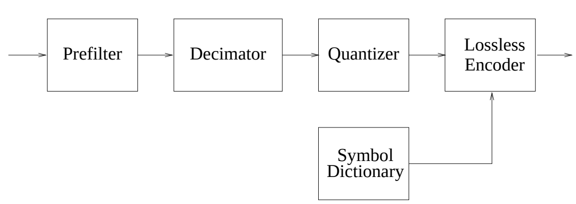


Figure 1: Proposed JBIG2 encoding method for compressing halftones. Free parameters are pattern grid size M, number of quantization levels N, and sharpening control L. The prefilter would be applied only for stochastic halftones.

4. QUALITY METRICS

We develop quality metrics to evaluate the performance of the proposed encoding method. The proposed encoding method attempts to preserve the useful information present in the stochastic halftone while discarding as much noise and distortion as possible. So, we compare the input halftone with the original grayscale image.

Signal to noise ratios (SNRs) assume that the only source of degradation in a processed signal is additive noise. Compression and halftoning also introduce linear and nonlinear distortion. Because the human visual system responds independently to linear distortion and noise, we develop a quality measure for each effect [13, 18]. We quantify the linear distortion by constructing a minimum mean squared error Weiner filter so that the residual image is uncorrelated with the input image. The residual image represents the nonlinear distortion plus additive independent noise.

To quantify the effect of nonlinear distortion and noise on quality, we spectrally weight the residual by a contrast sensitivity function (CSF). A CSF is a linear approximation of the human visual system response to a sine wave of a single frequency [13, 18]. A lowpass CSF assumes that the observer does not focus on one point in the image but freely moves the eyes around the image. We form a weighted SNR

$$\mathrm{WSNR} = 10 \log_{10} \left(\frac{\sum_u \sum_v |X(u,v)C(u,v)|^2}{\sum_u \sum_v |D(u,v)C(u,v)|^2} \right)$$

where $C(u,v)$ is a lowpass CSF, and $X(u,v)$ and $D(u,v)$ are the Fourier transforms of windowed original and residual images, respectively. A higher WSNR means higher quality. To prevent WSNRs to be biased by large DC components, we initially remove the DC component of the images.

To quantify the effect of linear distortion, we compute a Linear Distortion Measure

$$\mathrm{LDM} = \frac{\sum_u \sum_v |1 - H(u,v)||X(u,v)C(u,v)|}{\sum_u \sum_v |X(u,v)C(u,v)|}$$

where $H(u,v)$ is the Fourier transform of the Weiner filter that models the process. $H(u,v)$ is lowpass, so $1 - H(u,v)$ is highpass. Because of the weighting by $X(u,v)$, the LDM only measures those frequencies which are present in the original image. A higher LDM means lower quality.

5. RESULTS

We compress the 512×512 Floyd-Steinberg error diffused halftone of the grayscale image barbara shown in Fig. 2. Figs. 3–9 show the effect of the encoding parameters. Table 1 gives the compression rates and distortion. The distortion measures assume a 600 dpi rendering device and a 40 cm viewing distance. For lossy compression, the encoder uses a pattern dictionary generated by clustered dot halftones [5].

As shown in Table 1, the existing Group 4 MMR fax standard expands the original image by 148% whereas the arithmetic JBIG2 generic MQ lossless coder gives a compression ratio of 1.73. If the image is encoded without a prefilter, then the noise introduced by stochastic halftoning is visible (Fig. 3). By prefiltering, we reduce the high-frequency noise and some of the detail (Fig. 4). We compensate for the frequency distortion by adjusting the sharpening parameter (Fig. 5). The sharpening parameter improves visual quality but decreases the compression ratio

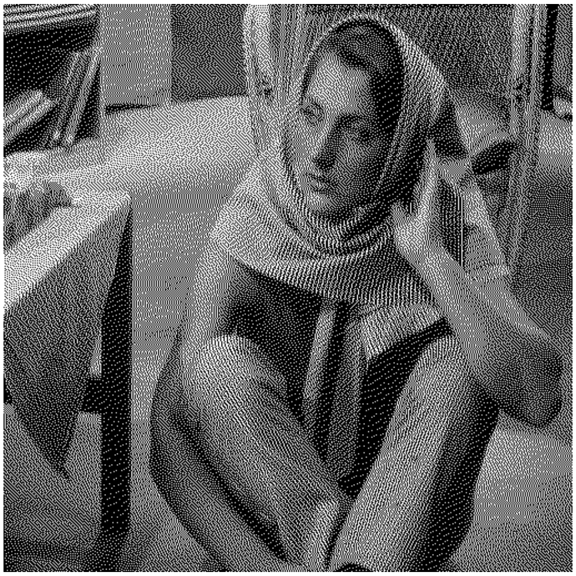


Figure 2: Original halftone of the barbara image.

Figure 3: Reconstructed halftone without prefiltering using a 4×4 grid ($M = 4$) and $N = 17$ gray levels. Same as [5].

(Fig. 4–6). Fig. 6 depicts oversharpening (contrast distortion).

Using a larger grid size and an angled screen can achieve better quality for the same bit rate (Figs. 5 and 7). Increasing the grid size increases compression (Figs. 7 and 8). Quantizing the descreened image to fewer gray levels further increases compression (Fig. 9). Rate-distortion curves (Fig. 10) represent the variation of distortion and compressed image size with respect to the grid size and sharpening control parameter. For linear distortion, an optimal value of the sharpening control parameter exists. Beyond the optimal value, the reconstructed halftones are oversharpened and distortion increases, which increases both rate and distortion. WSNR does not change significantly with the sharpening control parameter but changes dramatically with the grid size. The sharpness control parameter can be used to trade off compressed image size for improved quality.

	Encoder Parameters				Distortion		Size
Image	L	M	N	θ	*LDM*	*WSNR*	*Bytes*
Orig.	–	–	–	–	–	–	32768
MMR	–	–	–	–	–	–	81420
MQ	–	–	–	–	–	–	18907
Fig. 3	–	4	17	0^o	0.163	15.4 dB	5374
Fig. 4	–	4	17	0^o	0.181	16.5 dB	4356
Fig. 5	0.5	4	17	0^o	0.091	16.0 dB	5152
Fig. 6	1.5	4	17	0^o	0.292	14.8 dB	6352
Fig. 7	0.5	6	19	45^o	0.116	18.7 dB	4961
Fig. 8	0.5	8	33	45^o	0.155	15.7 dB	3983
Fig. 9	0.5	8	16	45^o	0.158	14.0 dB	3318

Table 1: Visual quality and size of the JBIG2 bitstream for the barbara image for the encoder parameters: sharpening control L, $M \times M$ block size, N quantization levels, and orientation θ. Section 4 defines LDM and WSNR. MMR and MQ are lossless JBIG2 modes. MMR is same as the current fax standard. MQ uses arithmetic coding.

Figure 4: Reconstructed halftone with a prefilter using a 4×4 grid ($M = 4$) with all 17 gray levels ($N = 17$).

6. CONCLUSION

We have developed a JBIG2-compliant method for encoding stochastic halftones. To descreen the halftone, we prefilter, decimate, and quantize. Prefiltering reduces high-frequency noise, spurious tones, and Nyquist frequencies. Decimation reduces spatial resolution. Quantization uses modified Floyd-Steinberg error diffusion to shape the quantization error into the higher frequencies and control the sharpness. We analyze the rate-distortion tradeoffs for the encoder parameters: orientation of the halftone grid; decimation factor which also specifies the halftone grid; number of quantization levels; and sharpening control. To measure

Figure 5: Reconstructed halftone with prefiltering and sharpening ($L = 0.5$) using a 4×4 grid ($M = 4$) with all 17 gray levels ($N = 17$).

Figure 6: Reconstructed halftone with prefiltering and sharpening ($L = 1.5$) using a 4×4 grid ($M = 4$) with all 17 gray levels ($N = 17$).

Figure 7: Reconstructed halftone with prefiltering and sharpening ($L = 0.5$) using a 6×6 grid ($M = 6$) angled at 45^o with all 19 gray levels ($N = 19$).

Figure 8: Reconstructed halftone with prefiltering and sharpening ($L = 0.5$) using an 8×8 grid ($M = 8$) angled at 45^o with all 33 graylevels ($N = 33$).

distortion, we develop a quality measure for linear distortion and one for nonlinear distortion and additive noise. For the same level of distortion, our method can achieve an additional 35% compression over the method in [5], as seen by comparing the results Figs. 3 and 8 in Table 1.

7. REFERENCES

[1] R. Ulichney, *Digital Halftoning.* MIT Press, 1987.

[2] *Fascimile Coding Schemes*, vol. VII. Geneva, Switzerland: CCITT Rec. T.4 T.6, 1984.

[3] K. Denecker, S. Van Assche, and I. Lemahieu, "A fast autocorrelation based context template selection scheme for lossless compression of halftone images," *Proc. SPIE Color Imaging Conf.*, vol. 3300, pp. pp. 262–272, Jan. 1998.

[4] D. A. D. Tompkins and F. Kossentini, "A fast segmentation algorithm for bi-level image compression using JBIG2," in *Proc. IEEE Int. Conf. Image Proc.*, Oct. 1999.

[5] B. Martins and S. Forchhammer, "Halftone coding with JBIG2," *IEEE Trans. Image Proc.*, to appear.

[6] T. D. Kite, N. Damera-Venkata, B. L. Evans, and A. C. Bovik, "A high quality, fast inverse halftoning algorithm for error diffused halftones," in *Proc. IEEE Int. Conf. Image Proc.*, vol. 2, pp. 64–8, Oct. 1998.

[7] N. Damera-Venkata, T. D. Kite, M. Venkataraman, and B. L. Evans, "Fast blind inverse halftoning," in *Proc. IEEE Int. Conf. Image Proc.*, vol. 2, pp. 71–4, Oct. 1998.

[8] J. Luo, R. de Queiroz, and Z. Fan, "A robust technique for image descreening based on the wavelet transform," *IEEE Trans. Signal Proc.*, vol. 46, pp. 1179–94, Apr. 1998.

[9] Z. Xiong, M. Orchard, and K. Ramachandran, "Inverse halftoning using wavelets," in *Proc. IEEE Int. Conf. Image Proc.*, vol. 1, pp. 569–72, Sept. 1996.

[10] N. T. Thao, "Set theoretic inverse halftoning," *Proc. IEEE Int. Conf. Image Proc.*, vol. 1, pp. 783–6, Oct. 1997.

[11] S. Hein and A. Zakhor, "Halftone to continuous-tone conversion of error-diffusion coded images," *IEEE Trans. Image Proc.*, vol. 4, pp. 208–16, Feb. 1995.

[12] J. Z. C. Lai and J. Y. Yen, "Inverse error-diffusion using classified vector quantization," *IEEE Trans. Image Proc.*, vol. 7, pp. 1753–8, Dec. 1998.

[13] T. D. Kite, *Design and Quality Assessment of Forward and Inverse Error Diffusion Halftoning Algorithms.* PhD thesis, The University of Texas, Austin, Texas, Aug. 1998.

[14] R. W. G. Hunt, *The Reproduction of Colour.* Fountain Press, 1995.

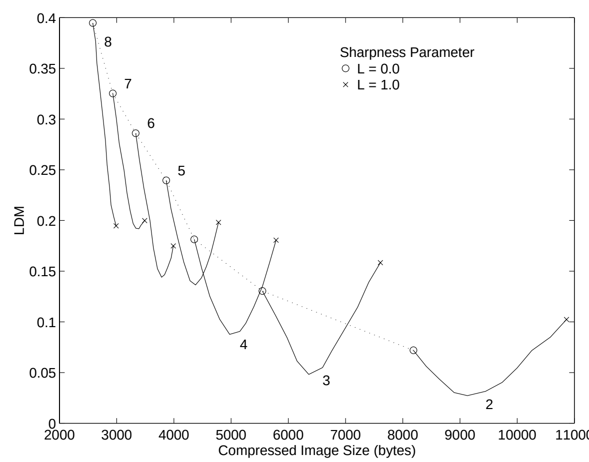
[15] T. Morita and H. Ochi, "Progressive transmission of continuous tone images using multi-level error diffusion method," *IEICE Trans. Comm.*, vol. 82-B, pp. 103–11, Jan. 1999.

[16] R. Eschbach and K. Knox, "Error-diffusion algorithm with edge enhancement," *J. Opt. Soc. Am. A*, vol. 8, pp. 1844–50, Dec. 1991.

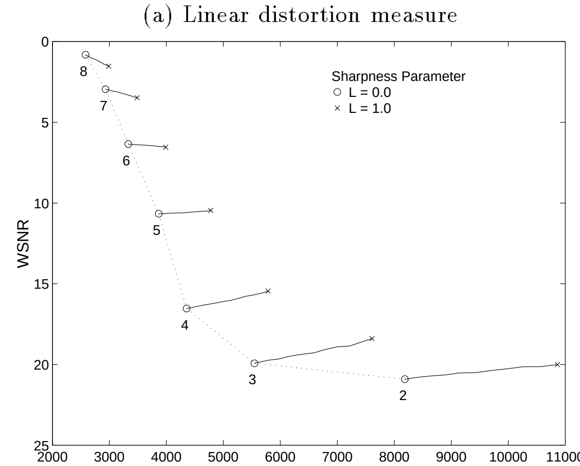
[17] R. Floyd and L. Steinberg, "An adaptive algorithm for spatial grayscale," *Proc. Soc. Image Display*, vol. 17, no. 2, pp. 75–7, 1976.

[18] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik, "Image quality assessment based on a degradation model," *IEEE Trans. Image Proc.*, to appear.

Figure 9: Reconstructed halftone with prefiltering, sharpening (L=0.5) and quantized to $N = 16$ gray levels using an 8×8 grid ($M = 8$) angled at 45^o.



(a) Linear distortion measure



(b) Weighted signal-to-noise ratio

Figure 10: Rate-distortion curves for the barbara image. Each curve is plotted by varying the sharpness parameter $L \in [0, 1]$ for a specific value of $M \in \{2, 3, 4, 5, 6, 7, 8\}$.